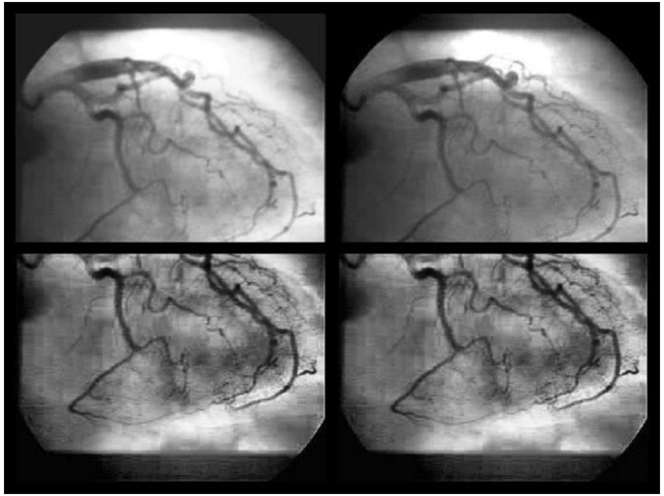
Numerical Algorithms for Solving Nonsmooth Optimization Problems and Applications in Image Reconstructions

Karina Rodriguez
(Joint work with Lewis Hicks, Mike Wells, and Nam Nguyen)

Department of Mathematics and Statistics
Portland State University

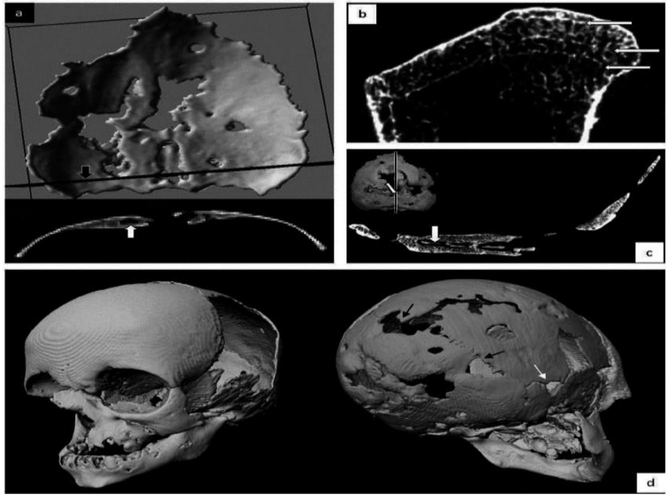
Image Reconstruction

3-D Reconstruction of the Heart



(2018, December 12). 3D reconstruction of the heart by advanced image processing algorithms.RSIP. Retrieved from https://www.rsipvision.com/3d-reconstruction-of-the-heart/

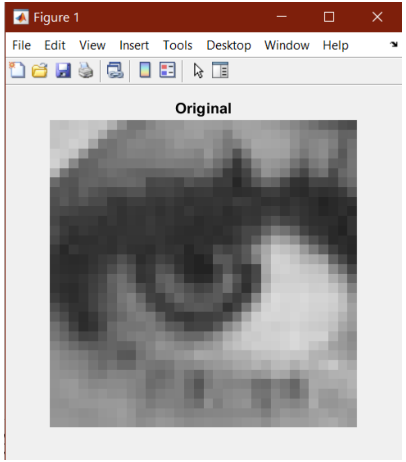
3-D Reconstruction of Skull



Colombo, Antony & Saint-Pierre, Christophe & Naji, Stephan & Panuel, Michel & Coqueugniot, Hélène & Dutour, Olivier. (2015). Langerhans cell histiocytosis or Tuberculosis on a medieval child (Oppidum de la Granède, Millau, France - 10th-11th centuries AD). Tuberculosis. 95. S42-S50. 10.1016/j.tube.2015.02.003.

Related Work and Our Approach

Minárová, Mária & Paternain, Daniel & Jurio, Aranzazu & Ruiz Aranguren, Javier & Takáč, Zdenko & Sola, Humberto. (2017). Modifying the Gravitational Search Algorithm: a functional study. Information Sciences. 430. 10.1016/j.ins.2017.11.033.



patching

NSF

REU

PICS PORTLAND INSTITUTE FOR COMPUTATIONAL SCIENCE

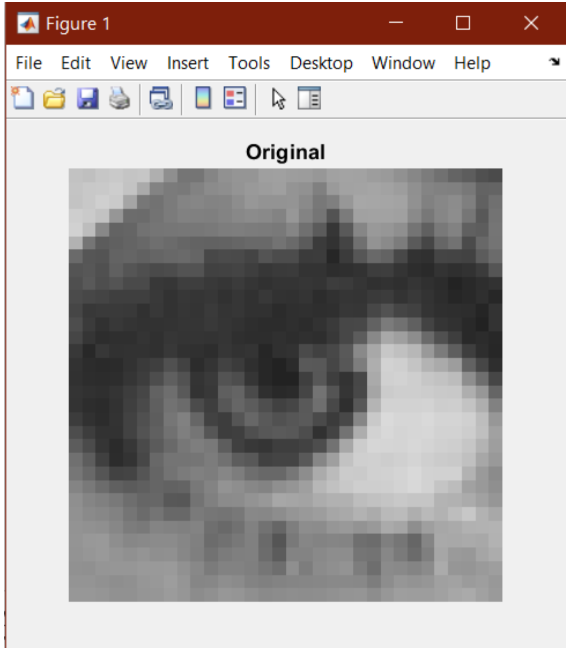
Portland State UNIVERSITY

Image in Code

80	88	97	102	102	99	96	90
80	97	105	104	99	96	95	89
84	101	108	103	95	93	93	91
91	101	103	98	92	91	92	95
99	99	96	91	89	90	91	100
04	97	90	87	87	89	91	103
98	90	84	82	84	89	92	99
93	86	81	80	85	92	96	107
86	80	76	79	87	97	103	115
78	75	74	80	90	102	109	116
73	74	78	85	96	106	112	109
71	78	86	94	102	108	112	101
71	83	95	104	108	109	109	98
72	87	101	110	112	110	107	99
77	96	107	116	120	113	101	111

Minárová, Mária & Paternain, Daniel & Jurio, Aranzazu & Ruiz Aranguren, Javier & Takáč, Zdenko & Sola, Humberto. (2017). Modifying the Gravitational Search Algorithm: a functional study. Information Sciences. 430. 10.1016/j.ins.2017.11.033.

Vectorizing Small Image



patch

$$\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{bmatrix} \rightarrow \begin{bmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \\ \vdots \\ x_{1,n} \\ \vdots \\ x_{n,n} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix}$$

vectorization

PICS PORTLAND INSTITUTE FOR COMPUTATIONAL SCIENCE

Portland State UNIVERSITY

Create Blurry Image

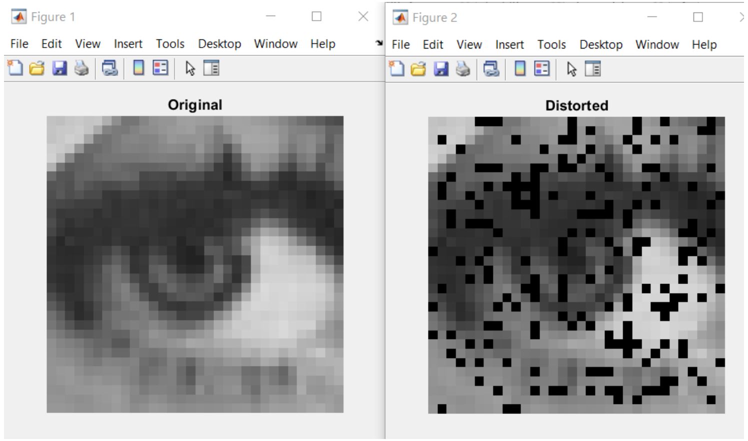
```
%extract patch size rows:251-282, col:251-282
%X is considered as the image
patch = X(251:282,251:282);
figure
imshow(patch) %show patch
truesize([300 300]);
title('Original'); %add title
sampleRate = .5;
%returns array [1 204] with random numbers 1 thru 64
Omega = randi([1,32*32], 1,floor(32*32/5));
%vector of numbers 1-1024
index = 1:1024;
%assign 0 (makes pixel black) to random (Omega) indices of index vector
index(Omega) = 0;
%gets indices ~0
pos = index(index > 0);
%create a 1024*1024 w/ 0's matrix with 1's along diagonal
T = eye(32*32);
%make A matrix to be all zeros, excluding random indices greater then 0
%i.e. indices 3 and 6, a matrix 16*16 the values of row1 col3 && row2 col6
%are 1
A = T(pos,:);
%create vector from image patch
X3 = patch(:);
%create image with missing pixels
b = A*X3;
c = reshape(A'*b, size(patch));
```

$$Ax + E = b$$

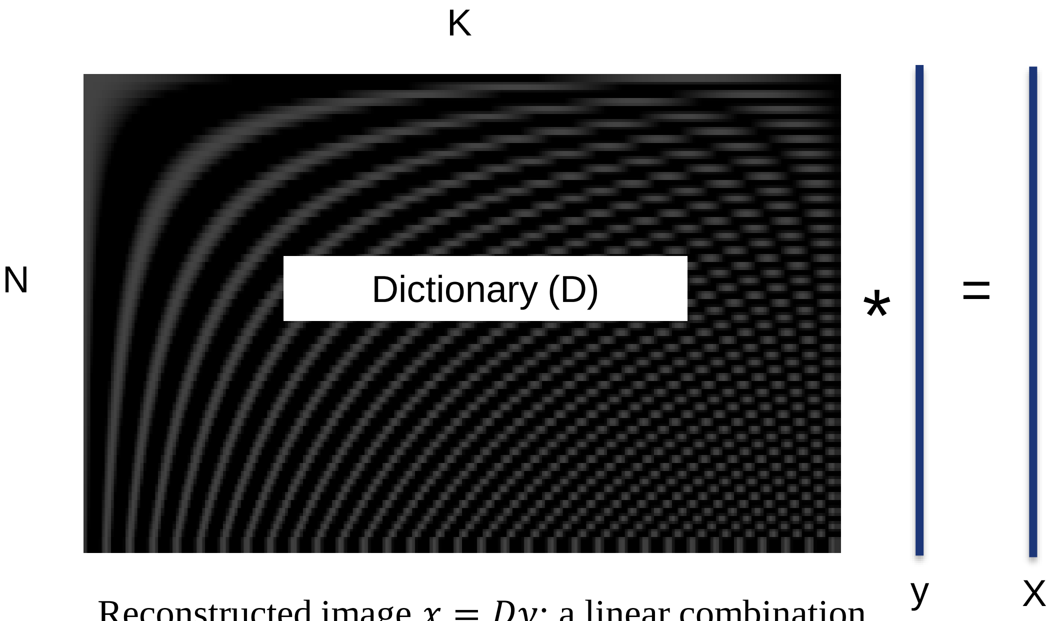
A deletes entries from the vectorized image x.

Identity matrix with missing rows

$$Ax = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$



Dictionary to Create a Good Image



Reconstructed image $x = Dy$: a linear combination of the columns (“atoms”) of D

Define $\mathcal{A} := AD$.

Properties of a “good image”:

- $Ax = A(Dy) = \mathcal{A}y$ is very close to b:

  $\frac{1}{2}\|\mathcal{A}y - b\|^2$ is small

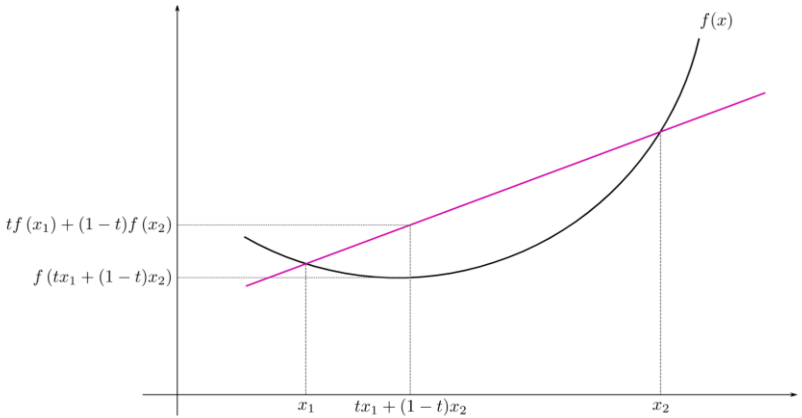
- The vector y is sparse:

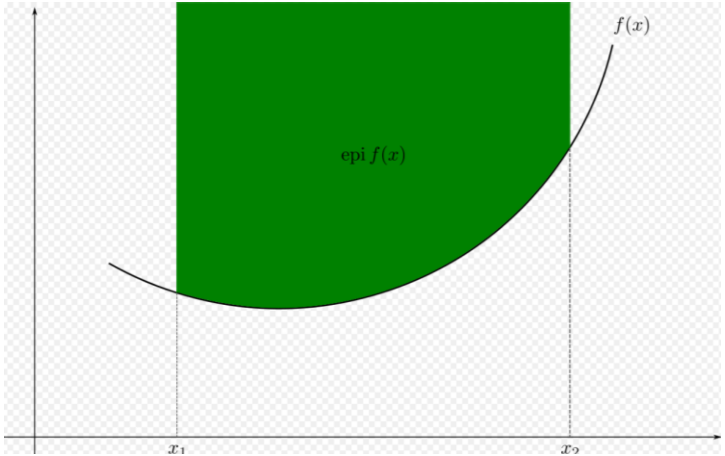
  $\|y\|_0 \approx \|y\|_1 - \|y\|_2$ is small

We seek to minimize

$$f(y) := \frac{1}{2}\|\mathcal{A}y - b\|^2 + \lambda(\|y\|_1 - \|y\|_2)$$

Convex Function



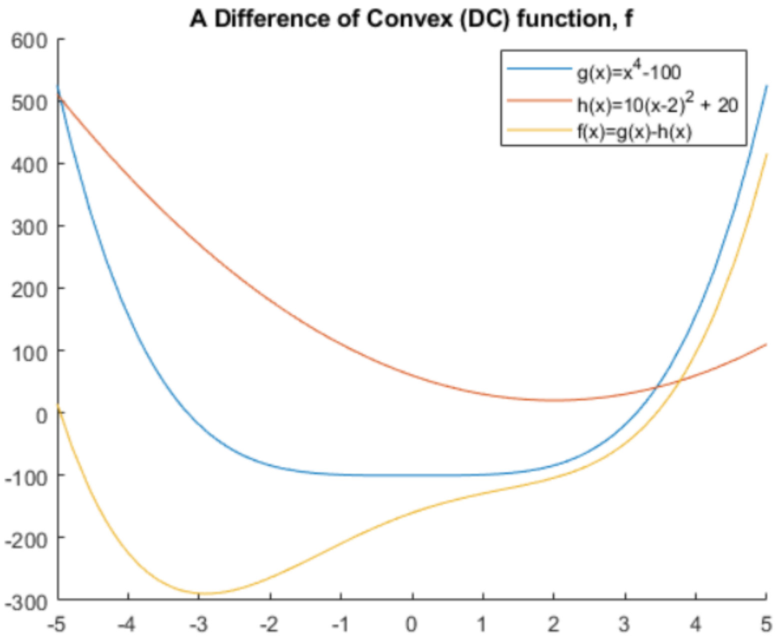


Wikipedia contributors. (2019, August 8). Convex function. In *Wikipedia, The Free Encyclopedia*. Retrieved 01:58, August 23, 2019, from https://en.wikipedia.org/w/index.php?title=Convex_function&oldid=909989312

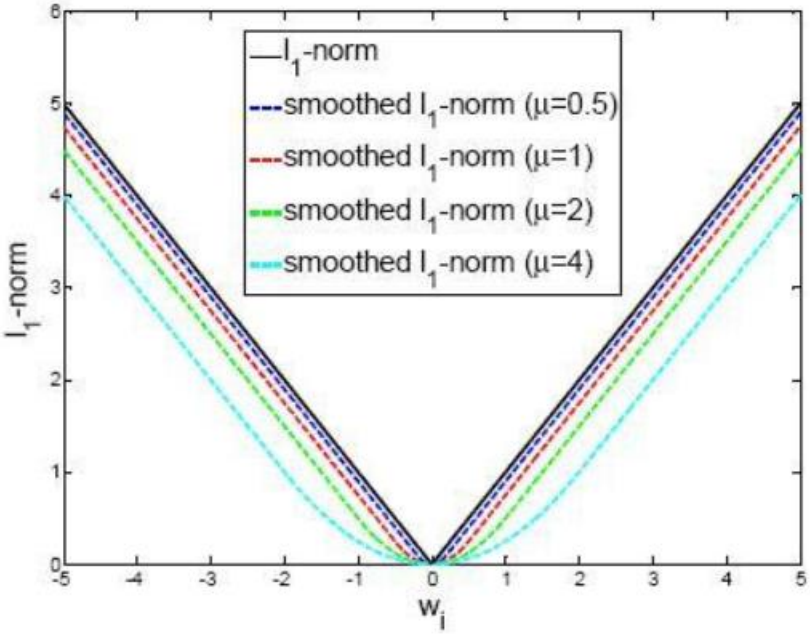
A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is called *convex* if for all $x_1, x_2 \in \mathbb{R}^n$ and for all t ∈ [0,1]

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

DC Functions



Nesterov Smoothing



Equation being smoothed:

$$y = |x|$$

DCA Algorithm

```
x=0*AT(b)+.5;
f=@(x) v/2*norm(A(x)-b)^2+norm(x,1);
for LP=1:N
    for i=1:3
        proj=max(-1, min(x/mu,1));
        if norm(x)==0
            w=0*x;
        else
            w=x/norm(x);
        end

        arg=x/mu;
        proj=max(-1, min(arg,1));
        y=w+gam*x-v*AT(A(x)-b)+(x/mu-proj);
        x=mu/(1+gam*mu)*y;

        ct=ct+1;
        FPLOT(ct)=f(x);
    end
    mu=mu*sig;
end
```

Algorithm 1 The DCA

1: **Input:** $x_0 \in \mathbb{R}^n$, $N \in \mathbb{N}$.
2: **for** $k = 1, \ldots, N$ **do**
3: Find $y_k \in \partial h(x_{k-1})$
4: Find $x_k \in \partial g^*(y_k)$
5: **end for**
6: **Output:** x_N.

Function: $f = \frac{1}{2}\|Ax - b\|^2 + \lambda\|x\|_1 - \lambda\|x\|_2$

which has

$$g = \frac{1}{2}\|Ax - b\|^2 + \lambda\|x\|_1$$

and $$h = \lambda\|x\|_2.$$

DCA with Smoothing

Since $\|x\|_1$ is non-smooth, we use Nesterov's Smoothing Technique to find a smooth approximation:

$$p_\mu(x) = \frac{1}{2\mu}\|x\|^2 - \frac{\mu}{2}d(\frac{x}{\mu}; Q)^2$$

where $Q = \{x \in \mathbb{R}^n \mid |x_i| \le 1\}$. The function

$$f(x) = \frac{1}{2}\|Ax - b\|_2^2 + \lambda(\|x\|_1 - \|x\|_2)$$

is then approximated by

$$f_\mu(x) = \frac{\lambda}{2\mu}\|x\|^2 + \frac{1}{2}\|Ax - b\|^2 - \frac{\lambda\mu}{2}d(\mu^{-1}x; Q)^2 - \lambda\|x\|.$$

DCA with Smoothing

We set

$$g(x) = \left(\frac{\lambda + \gamma\mu}{2\mu}\right) \|x\|^2$$

and

$$h(x) = \frac{\gamma}{2}\|x\|^2 - \frac{1}{2}\|Ax - b\|^2 + \frac{\lambda\mu}{2} d(\mu^{-1}x; Q)^2 + \lambda\|x\|.$$

Gamma is chosen large enough so $\frac{\gamma}{2}\|x\|^2 - \|Ax - b\|^2$ is convex.

We have

$$\partial h(x) = \left(\frac{\lambda + \gamma\mu}{\mu}\right) x - A^T(Ax - b) - \lambda\Pi_Q(\mu^{-1}x) + \lambda\partial\|x\|$$

where Π_Q is projection onto Q.

DCA with Smoothing

We substitute

$$\omega(x) = \begin{cases} \frac{x}{\|x\|} & x \neq 0, \\ 0 & x = 0, \end{cases}$$

for $\partial\|x\|$ since it is a subgradient for all x.

We use the fact that $x \in \partial g^*(y)$ if and only if $y \in \partial g(x)$. Thus,

$$y = \left(\frac{\lambda + \gamma\mu}{\mu}\right) x$$

implies

$$x = \left(\frac{\mu}{\lambda + \gamma\mu}\right) y.$$

The Boosted DCA

Boosted DCA Algorithm

INPUT: x_0, $N \in \mathbb{N}$,

$\alpha > 0, \bar{\lambda} > 0, 0 < \beta < 1.$

for $k = 0, \ldots, N$ **do**

1. Find $z_k \in \partial h(x_k)$.
2. Solve $y_k = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \{g(x) - \langle z_k, x \rangle\}$.
3. Set $d_k = y_k - x_k$. If $d_k = 0$, stop, return x_k. Else, continue.
4. Set $\lambda_k = \bar{\lambda}$.

   While $\phi(y_k + \lambda_k d_k) > \phi(y_k) - \alpha \lambda_k \|d_k\|^2$, set $\lambda_k = \beta \lambda_k$
5. Set $x_{k+1} = y_k + \lambda_k d_k$.

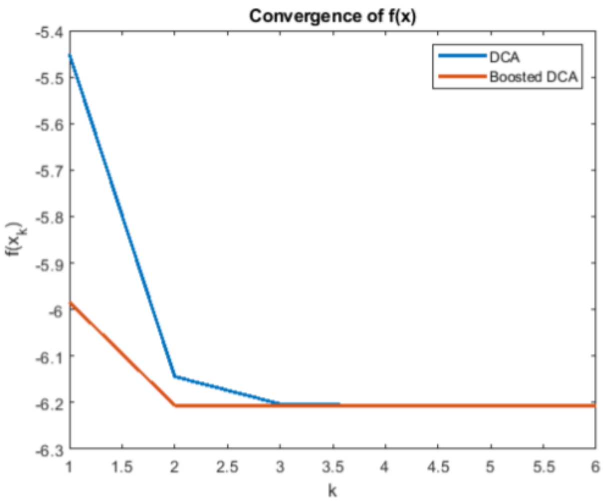
If $x_{k+1} = x_k$, stop, return x_k. Else, set $k = k + 1$ and return to step 1.

end for

OUTPUT: x_{N+1}

DCA and Boosted DCA

$$\text{minimize } f(x) = x^4 - 2x^2 + 2x - 3, \;\; x \in \mathbb{R}$$

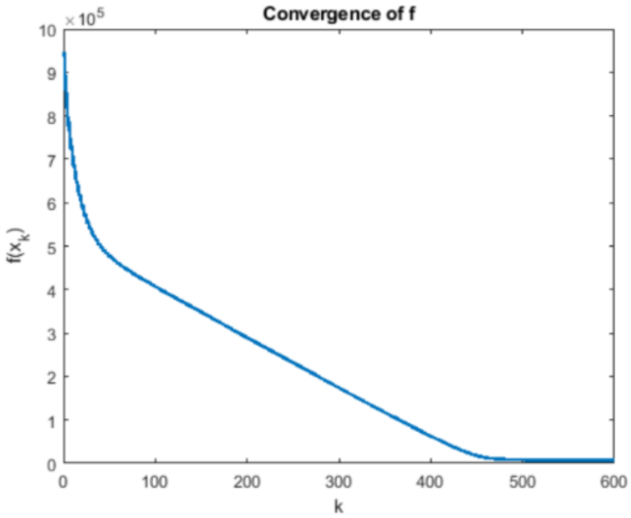


RE and PSNR

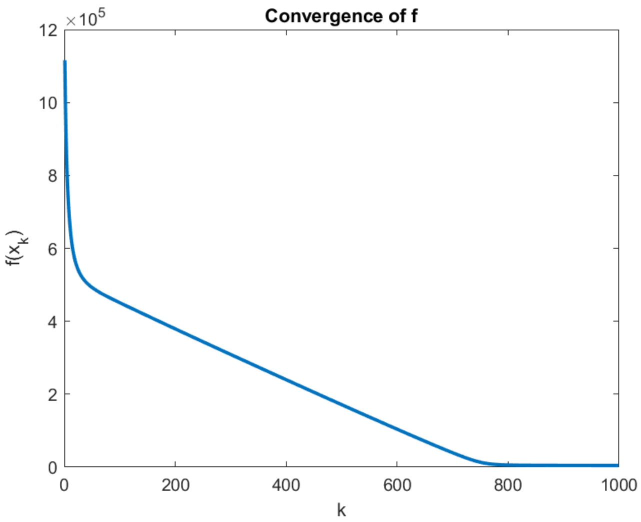
$$RE = 100\% * \frac{||M - \widehat{M}||}{||M||}$$

$$PSNR = 20log_{10}(\frac{\sqrt{N_1 N_2}}{||M - \widehat{M}||})$$

Convergence Graph



Boosted DCA with Nesterov Smoothing



DCA with Nesterov Smoothing

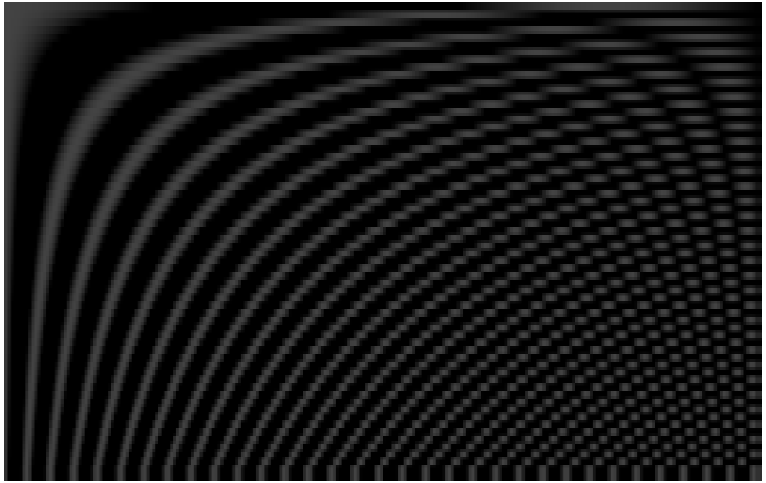
Results

Sampled Image

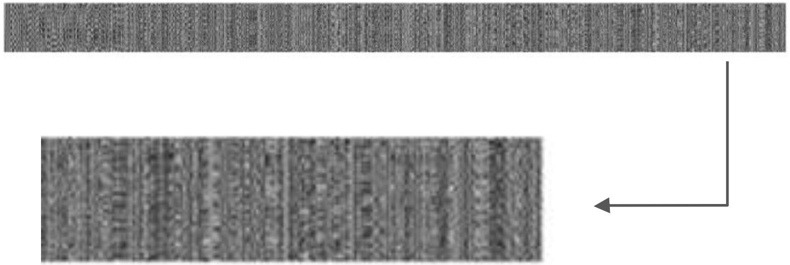
Boosted DCA

DCA

Future Work



DCT Dictionary



Learned Dictionary

Computational Modeling Serving the City

(2016, March 7). Retrieved from https://www.youtube.com/watch?v=qMhGzaPhTKk

Low-quality video

Acknowledgements

Lewis Hicks
Michael Wells
Advisor: Mau Nam Nguyen

The REU Site is supported by the National Science Foundation under grant no 1758006.

References

- N.M. Nam, L.T.H. An, N.T. An, D.Giles, Smoothing techniques and difference of convex functions algorithms for image reconstructions, Optimization (2019), accepted.
- Nesterov Y. Smooth minimization of non-smooth functions. Math.Program., Ser. A, 103 (2005), 127–152.
- Pham Dinh T, Le Thi HA. Convex analysis approach to D.C. programming: Theory, algorithms and applications. Acta Math. Vietnam. 22 (1997), 289–355.
- Pham Dinh T, Le Thi HA. A d.c. optimization algorithm for solving the trust-region subproblem. SIAM J. Optim., 8 (1998), 476–505.
- Vandenberghe L. Optimization methods for large-scale systems, EE236C lecture notes, UCLA.
- Xin J, Osher S, Lou Y. Computational aspects of L1-L2 minimization for compressive sensing. Advances in Intelligent Systems and Computing, 359 (2015), 169–180.
- Yin P, Lou Y, He Q, Xin J. Minimization of L1-L2 for compressed sensing. SIAM J. of Sci. Comput. 37 (2015), A536–A563.